Fade and Dissolve Detection in Uncompressed and Compressed Video Sequences

W.A.C. Fernando, C.N. Canagarajah, D. R. Bull

Image Communications Group, Centre for Communications Research, University of Bristol,
Merchant Ventures Building, Woodland Road, Bristol BS8 1UB, United Kingdom.
Voice - +44-117-954-5198, Fax - +44-117-954-5206
Email - W.A.C.Fernando@bristol.ac.uk

Abstract

Automatic identification of special effects is a pre-requisite for video indexing and intelligent video encoding. In this paper we present an algorithm for fade and dissolve scene change detection in video sequences. We use statistical features of the images to identify these special effects in uncompressed video. DC-estimation [4] is used to evaluate statistical features both in H.263 [10] and MPEG-2 [11] compressed video. Results show that these special effects can be identified accurately with the proposed scheme.

1. INTRODUCTION

Due to rapid advances in compression technology and imaging hardware, the expansion of low-cost storage media, and the explosion of the internet, the digital video "for every one" is now becoming a possible target. The demand for digital video is also increasing in areas such as video conferencing, multimedia authoring systems, education and video-on-demand systems. A major feature required in a visual information system is an efficient indexing to enable fast access to the stored data. A common and natural idea is to index the video sequences first into video shots by identifying scene changes and then to extract features. Therefore a powerful scene change detection, which can operate both in compressed and uncompressed domain, is required to allow for a complete characterisation of the video sequences.

Abrupt transitions are very easy to detect as the two frames are completely uncorrelated. But, gradual transitions are more difficult to detect as the difference between frames corresponding to two successive shots is substantially reduced. Considerable work has been reported on detecting abrupt transitions both in uncompressed video and compressed video [1-5]. However, only a small effort has been directed toward gradual scene change detection [1,4,6-9]. Zabith et al [1] proposed a feature-based algorithm for detecting and classifying scene beaks. This algorithm needs to detect edges in every frame and hence it is very costly. Another limitation of this scheme is that the edge detection method does not handle rapid changes in overall scene brightness, or scenes, which are very dark or very bright. Furthermore, automatic segmentation and classification is not possible with this scheme. Alattar proposed an algorithm for detecting fade-in and fade-out by exploiting the semi-parabolic behaviour of the variance curve [8]. This algorithm can only detect fade-in and fade-out when the end frames are fixed (freezed). When the sequence has considerable motion, this algorithm fails to identify fade-in and fade-out regions.

In this paper we present an algorithm for dissolving and fading scene change detection in video sequences. We exploit statistical features of the luminance signal of the video to detect these gradual transitions. Rest of the paper is organised as follows: Section 2 presents a mathematical model for fading and dissolving operations in video production. Section 3 describes the proposed algorithm for fade and dissolve detection. Results are presented in section 4. Section 5 discusses the conclusions and future work.

2. MATHEMATICAL MODEL FOR DISSOLVING AND FADING

In video editing and production, proportions of two or more picture signals are simply added together so that the two pictures appear to merge on the output screen. Very often this process is used to move on from picture A to picture B. In this case, the proportions of the two signals are so that as the contribution of picture A changes from 100% to zero and the contribution of picture B changes from zero to 100%. This is called dissolving. When picture A is a solid colour, it is called as fade-in and when picture B is a solid colour, it is known as fade-out. Mathematically, dissolving can be expressed as follows.

$$S_n(x,y)=\begin{cases} f_n(x,y) & 0\le n< L_1 \\ \left[1-\left(\frac{n-L_1}{F}\right)\right]f_n(x,y)+\left[\frac{n-L_1}{F}\right]g_n(x,y) & L_1\le n\le (L_1+F) \\ g_n(x,y) & (L_1+F)<n\le L_2 \end{cases} \quad (1)$$

0-7803-5467-2/99/ $ 10.00 © 1999 IEEE

Where, $S_n(x,y)$ - Resultant video signal, $f_n(x,y)$ - Picture A, $g_n(x,y)$ - Picture B, L_1 - Length of sequence A, F - Length of dissolving sequence, L_2 - Length of the total sequence.

Assume the video sequences $f_n(x,y)$ and $g_n(x,y)$ are ergodic processes with mean m_f, m_g and variance of σ_f^2, σ_g^2 respectively. Let, $m_{s,n}$ is the mean of the resultant video sequence (S_n) and $\sigma_{s,n}^2$ is the variance of the resultant video sequence. Equation (2) and (3) show the behaviour of mean and variance of the dissolved sequence.

$$m_{s,n} = E[S_n(x,y)]$$

$$= \begin{cases} m_f & 0 \le n < L_1 \\ \left[m_f - \frac{L_1}{F}(m_g - m_f)\right] - \frac{n}{F}(m_f - m_g) & L_1 \le n \le (L_1+F) \\ m_g & (L_1+F) < n \le L_2 \end{cases} \tag{2}$$

$$\sigma_{s,n}^2 = E[S_n^2] - E[S_n]^2$$

$$= \begin{cases} \sigma_f^2 & 0 \le n < L_1 \\ \Phi & L_1 \le n \le (L_1+F) \\ \sigma_g^2 & (L_1+F) < n \le L_2 \end{cases} \tag{3}$$

where,

$$\Phi = \left[\frac{\sigma_f^2 + \sigma_g^2}{F^2}\right] n^2 - \left[\frac{2L_1(\sigma_f^2+\sigma_g^2)}{F^2} - \frac{2\sigma_f^2}{F}\right] n + \left[\sigma_f^2 + \frac{L_1^2(\sigma_f^2+\sigma_g^2)}{F^2} + \frac{2L_1\sigma_f^2}{F}\right]$$

Similar arguments can be followed for fade-in and fade-out by simplifying the above equations. For example, fade-in can be described with the initiative function as shown in Equation (4). Equation 5 and 6 present the behaviour of mean and variance respectively.

$$S_n(x,y) = \begin{cases} f_n(x,y) & 0 \le n < L_1 \\ \left[1 - \left(\frac{n-L_1}{F}\right)\right] C + \left[\frac{n-L_1}{F}\right] g_n(x,y) & L_1 \le n \le (L_1+F) \\ g_n(x,y) & (L_1+F) < n \le (L_2) \end{cases} \tag{4}$$

where C is a solid colour.

$$m_{s,n} = \begin{cases} m_f & 0 \le n < L_1 \\ \left[C + \frac{L_1}{F}(C - m_g)\right] + \frac{n}{F}(m_g - C) & L_1 \le n \le (L_1+F) \\ m_g & (L_1+F) < n \le L_2 \end{cases} \tag{5}$$

$$\sigma_{s,n}^2 = \begin{cases} \sigma_f^2 & 0 \le n < L_1 \\ \Phi & L_1 \le n < (L_1+F) \\ \sigma_g^2 & (L_1+F) \le n < L_2 \end{cases} \tag{6}$$

where, $\Phi = \left[\frac{\sigma_g^2}{F^2}\right]\left[n^2 - 2nL_1 + L_1^2\right]$

3. FADE AND DISSOLVE DETECTION

Thus, it is clear that during fading/dissolving, mean and variance have a linear and a quadratic behaviour respectively. However, these mathematical are valid under the assumption that the video sequences are an ergodic process. In practice, an ergodic process cannot be guaranteed always due to motion in video. Therefore, these statistical behaviours may be slightly distorted for a practical video sequence. Thus, some alternative strategies are needed along with mean and variance of the video sequences in order to identify these special effects.

3.1 Detection of Fading and Dissolving in Uncompressed Video

We can approximate that the mean and the variance have a linear and quadratic behaviour during fading/dissolving. However, it is not possible to identify these special effects by considering either mean or variance individually. Therefore, in this proposed scheme we combine both these features in order to identify fading/dissolving. Since mean has a linear behaviour the first derivative should be a constant during the dissolving period. Second derivative of the variance should also be a constant as variance curve has a quadratic behaviour during a dissolving period. Therefore, the ratio between the second derivative of the variance curve and the first derivative of the mean curve should be a constant. The differentiation of this ratio is checked to identify dissolving with a threshold T_{dis}. There may be a possibility of satisfying this condition for a small number of consecutive frames in a non-dissolve sequence. But, small sizes of dissolve sequences are not common in practice. Using this argument false regions are eliminated in the proposed algorithm. Furthermore, if there are two consecutive dissolve regions separated by very small gap, they are bridged to form a longer dissolve region.

This argument is extended to detect fade-in and fade-out since fading is a special case of dissolving, where one scene is a solid colour. It should be noted that variance is zero at the start of a fade-in sequence and at the end of a fade-out sequence. Finally, fade-in, fade-out and dissolving are identified as follows.

Fade-in Detecting a frame with zero variance followed by a sequence of continuos region, which is identified by the threshold T_{dis}.

Fade-out Detecting a sequence of continuos region, which is identified by the threshold T_{dis} followed by a frame with zero variance.

Dissolving Detecting a sequence of continuos region, which is identified by the threshold T_{dis}.

3.2 Detection of Fading and Dissolving in Compressed Video

Algorithms, which are proposed for uncompressed video, can still be applied in compressed domain by considering DC-sequences [4] of each frame. In compressed domain, it is considered 8x8 macroblocks (MBs) and all DC values of each MB are calculated using DC-sequence scheme [4]. Using this, mean of each frame can be evaluated approximately. However, variance may be slightly distorted due to the assumption of all the pixels have the same value in a particular MB.

4. RESULTS

4.1 Fading

Figure 1 and Figure 2 show the mean and variance respectively for the first test video sequence. This test sequence contains one fade-in and one fade-out gradual transitions as shown in Figures 1-2. Figure 3 shows the absolute values of the differentiation in ratio between the second derivative of the variance curve to the first derivative of the mean curve. There are two regions (Figure 3), which identifies by the algorithm with the threshold T_{dis}=2, right after the 31st frame and 111th frame. Considering the variance of the sequence, we can distinguish fade-in and fade-out regions as discussed in section 3. Therefore, fade-in region and fade-out regions are identified from 31st frame to 60th frame and from 111th frame to 150th frame respectively.

We considered a video sequence of 2200 frames to test the proposed algorithm. This sequence contains several other special effects such as wiping, panning and dissolving. Table 1 shows the summarised results with the proposed algorithm. Results show that the algorithm is capable of detecting all fade regions accurately even when the video sequence contains other special effects. Therefore, the proposed algorithm can be used in uncompressed video to detect fade regions with a high reliability.

Actual fade region	Detected fade region	Nature of the region
31-60	31-60	Fade-in
111-150	111-150	fade-out
248-303	249-303	fade-out
576-624	576-625	Fade-in
754-778	754-778	fade-out
944-986	944-987	Fade-in
1102-1167	1102-1168	Fade-in
1365-1420	1366-1420	fade-out
1500-1550	1500-1550	Fade-in
1620-1680	1620-1681	Fade-in
1760-1840	1761-1840	fade-out
1920-1985	1920-1985	fade-out

Table 1: Fade region identification : uncompressed video

4.2 Dissolving

Figure 4 and 5 show the behaviour of mean and variance of the second test sequence considered. Figure 6 shows the absolute values of the differentiation in ratio between the second derivative of the variance curve to the first derivative of the mean curve. Thus, dissolve regions can easily be identified with the threshold T_{dis}=2. Hence, the dissolve regions can be identified as frame 31-60 and 121-180. We considered 1500 frames sequence with this algorithm and Table 2 shows the summarised results.

Actual dissolve region	Detected dissolve region
31-60	31-60
121-180	121-180
221-280	221-280
325-385	324-385
446-496	446-497
548-604	548-604
804-868	804-869
1010-1089	1010-1089
1168-1232	1169-1232
1356-1424	1356-1424

Table 2: Dissolve region identification : uncompressed video

4.3 Fading and Dissolving in Compressed Video

Table 3 and table 4 show the results for fading and dissolving with H.263 [10] and MPEG-2 [11] compressed video. Results are slightly distorted, as variance calculation is not accurate with DC-estimation. But results are still encouraging.

Actual fade region	Detected region (H.263)	Detected region (MPEG-2)	Nature of the region
31-60	31-60	31-60	fade-in
111-150	112-150	112-150	fade-out
248-303	251-303	250-303	fade-out
576-624	576-627	576-627	fade-in
754-778	755-778	754-778	fade-out
944-986	944-988	944-987	fade-in
1102-1167	1102-1169	1102-1169	fade-in
1365-1420	1367-1420	1367-1420	fade-out
1500-1550	1500-1552	1500-1552	fade-in
1620-1680	1620-1682	1620-1681	fade-in
1760-1840	1762-1840	1762-1840	fade-out
1920-1985	1921-1985	1920-1985	fade-out

Table 3: Fade region identification : compressed video with the same sequence in 4.1

Actual dissolve region	Detected dissolve region (H.263)	Detected dissolve region (MPEG-2)
31-60	32-60	32-60
121-180	123-180	122-180
221-280	221-282	221-281
325-385	324-387	324-387
446-496	447-497	446-497
548-604	549-605	548-605
804-868	804-869	804-868
1010-1089	1012-1089	1012-1089
1168-1232	1169-1233	1169-1233
1356-1424	1356-1424	1356-1424

Table 4: Dissolve region identification : compressed video with the same sequence in 4.2

5. CONCLUSIONS

In this paper, we presented an algorithm for fade and dissolve scene change detection in video sequences. Statistical features of each image have been used to identify these special effects. DC-estimation [4] is used to calculate mean and variance in both H.263 and MPEG-2 compressed video. Results for uncompressed video show that these special effects can be identified accurately with the proposed scheme. Compressed domain results are slightly distorted due to the limitations of variance calculation. Further work is required to evaluate compressed domain variance accurately in order to get better results with the proposed schemes.

ACKNOWLEDGEMENTS

First author would like to express his gratitude and sincere appreciation to the university of Bristol and CVCP for providing financial support for this work.

References

1. Zabith, R., Miller, J., and Mai, K., *"Feature-Based Algorithms for Detecting and Classifying Scene Breaks,"* 4th ACM Intl. Conference on Multimedia, San Francisco, California, November 1995.

2. Fernando, W.A.C., Canagarajah, C.N., Bull, D. R., *"Sudden Scene Change Detection in MPEG-2 Video Sequences,"* Paper Number - 13, Proceedings - International Workshop on Multimedia Signal Processing, 1999.

3. Fernando, W.A.C., Canagarajah, C.N., Bull, D. R., *"DFD Based Scene Segmentation For H.263 Video Sequences,"* Paper Number-747, Proceedings - IEEE International Symposium on Circuits and Systems (ISCAS), 1999.

4. Yeo, B.L., *"Rapid Scene Analysis on Compressed Video,"* IEEE transaction on circuits and systems for video technology, Vol. 5, No 6, pp. 533-544, 1995.

5. Shin, T. et. al, *"Hierarchical scene change detection in an MPEG-2 compressed video sequence"*, Proceedings - ISCAS, Vol.4, pp. 253-256, 1998.

6. Fernando, W.A.C., Canagarajah, C.N., Bull, D. R., *"Automatic Detection of Fade-in and Fade-out in Video Sequences,"* Paper Number-748, Proceedings - ISCAS, 1999.

7. Fernando, W.A.C., Canagarajah, C.N., Bull, D. R., *"Video Segmentation and Classification for Content Based Storage and Retrieval Using Motion Vectors,"* Storage and Retrieval for Image and Video Databases VII, SPIE, pp. 687-698, 1999.

8. Alattar, A.M., *" Detecting Fade Regions in Uncompressed Video Sequences,"* Proceedings of ICASSP'97, pp. 3025-3028, 1997.

9. Nagasaka, and Y. Tanaka, *"Automatic Video Indexing and Full-Video Search for Object Appearances,"* Visual Database Systems II, Eds.E.Kunth, and L.M. Wegner, Elsevier Science Publishers B.V., IFIP, pp. 113-127,1992.

10. ITU-T, Draft H.263, May 1996.

11. Thomas. S.,*"MPEG Digital Video-Coding standards,"* IEEE Signal Processing magazine, pp. 82-100, September-1997.

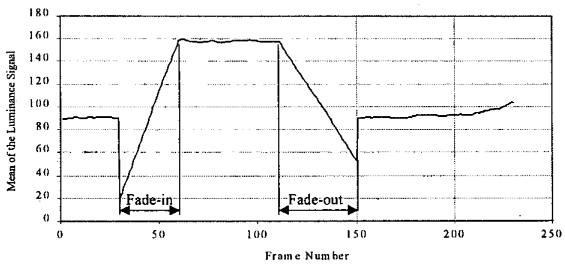

Figure 1: Mean of the Luminance signal (Fading Sequence)

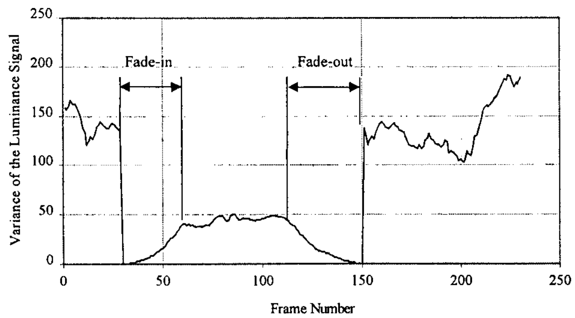

Figure 2: Variance of the Luminance signal (Fading Sequence)

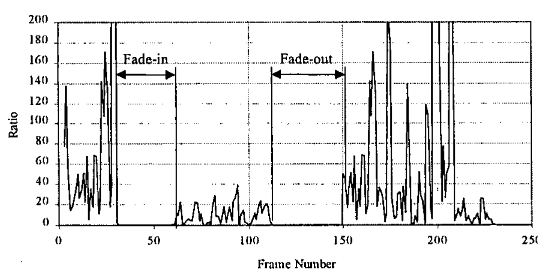

Figure 3: Absolute change in the ratio between second derivative of the variance curve to the first derivative of mean curve (Fading Sequence)

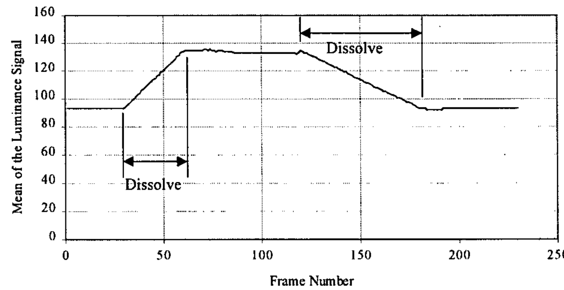

Figure 4: Mean of the Luminance signal (Dissolving Sequence)

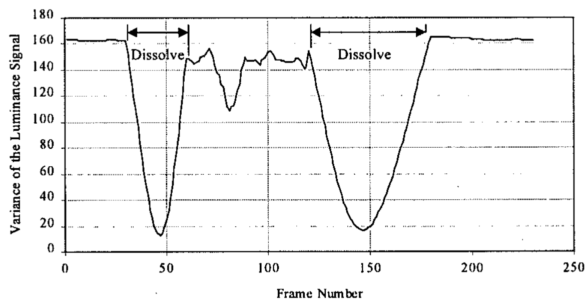

Figure 5: Variance of the Luminance signal (Dissolving Sequence)

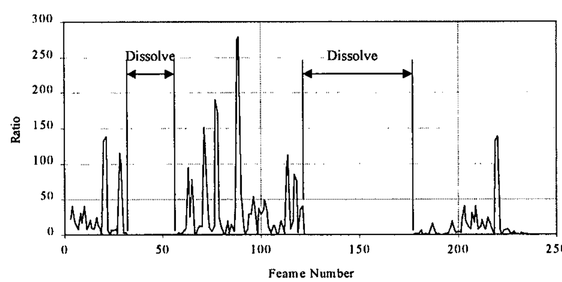

Figure 6: Absolute change in the ratio between second derivative of the variance curve to the first derivative of mean curve (Dissolving Sequence)